Stanford CS 229 Fall 2017

Final Project Milestone

Yang "Eddie" Chen, Yi Zhong, Yubo Tian

November 16, 2017

1. Motivation

Tennis matches are fun to watch because they are full of surprises. In the 2017 Stuttgart Open, Roger Federer, then an 18-time grand slam champion, lost to the world No. 302 player Tommy Haas. Federer lost the opening match at a grass-court tournament, a phenomenon that hadn't happened since 2002. How can we predict a rare loss like this? Based on past performance alone, any model would have predicted a Federer win in pre-game bets. This motivates us to apply machine learning to predict tennis matches in real time; in particular, we want to explore how well we can combine both historical performance data and real-time, just-happened set-by-set outcome, to achieve a better prediction. As a consequence, both tennis players and sports bettors can benefit from better predictions and new insights.

Tennis is an ideal candidate for a hierarchical model as a match consists of a sequence of sets, which in turn consist of a sequence of games, which in turn consist of a sequence of points...**TODO**: add more here to explain we start by looking at set-by-set

This paper seeks to model men's professional singles matches...**TODO**: *add some introduction after we have better ideas*

TODO: add literature review

2. Method

3. Preliminary Experiments

4. Next Steps

5. Contributions

5.1. Yi Zhong

5.2. Yubo Tian

5.3. Yang "Eddie" Chen

6. Conclusion

A. Appendix: Roadmap with Extensions

- **11/20 - 11/26**:
- **11/27 - 12/03**:
- **12/04 - 12/10**: Poster making
- **12/11 - 12/15**: Poster presentation; final project write-up